



## Festschrift in Honor of Dr. Robert Stickgold

### Original Article

# The influence of intentions on dream content

Julia Fechner<sup>1</sup>, Maren Born<sup>1</sup>, Massimiliano Mancini<sup>2</sup>, Zeynep Akata<sup>3,4</sup>, Philipp Haag<sup>1</sup>, Susanne Diekelmann<sup>1,5,†</sup> and Jan Born<sup>1,6,7,8,†,\*</sup>

<sup>1</sup>Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Tübingen, Germany,

<sup>2</sup>Department of Information Engineering and Computer Science, Multimedia and Human Understanding Group, University of Trento, Trento, Italy,

<sup>3</sup>Chair of Interpretable and Reliable Machine Learning, Technical University of Munich, Munich, Germany,

<sup>4</sup>Institute for Explainable Machine Learning, Helmholtz Munich, Munich, Germany,

<sup>5</sup>Department of Psychiatry and Psychotherapy, University Hospital Tübingen, Tübingen 72070, Germany,

<sup>6</sup>Werner Reichert Center for Integrative Neuroscience, University of Tübingen, Tübingen, Germany,

<sup>7</sup>German Center for Mental Health (DZPG), Tübingen, Germany and

<sup>8</sup>German Center for Diabetes Research (DZD), Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Center Munich at the University Tübingen (IDM), Tübingen, Germany

<sup>†</sup>Shared senior authorship.

\*Corresponding author: Jan Born, Institute for Medical Psychology and Behavioral Neurobiology, University of Tübingen, Otfried-Müller-Straße 25, 72076 Tübingen, Germany. Email: [jan.born@uni-tuebingen.de](mailto:jan.born@uni-tuebingen.de).

### Abstract

**Study Objectives:** The “Zeigarnik effect” refers to the phenomenon where future intentions are remembered effectively only as long as they are not executed. This study investigates whether these intentions, which remain active during sleep, influence dream content.

**Methods:** After an adaptation night, each of the 19 participants (10 women and 9 men) received three different task plans in the evening before the experimental night, each describing how to perform specific tasks. One of the task plans (completed) was then to be executed before the sleep period, another task (uncompleted) was told to be executed in the next morning, and on the third task (interrupted) participants were interrupted during the enactment before sleep and told to resume it the next morning. Polysomnography and multiple awakenings were conducted, resulting in 86 dream reports, 36 in NREM stage 2, and 50 in rapid eye movement sleep. After a traditional rating-based analysis of dream reports yielded inconsistent results, we analyzed the reports using a transformer-based assessment of dream incorporation, which quantified the semantic similarity between the dreams and pre-sleep tasks.

**Results:** The number of dreams showing above-criterion similarity to the respective task was significantly lower for the completed than the uncompleted or interrupted tasks ( $p < .05$ ,  $\chi^2$  test). This pattern was confirmed through a forced choice approach, where—based on the similarity of single sentences of the dream reports—each dream report was allocated to one of the three task plans ( $p < 0.01$ , one-tailed  $\chi^2$  test).

**Conclusions:** Active intentions increase the likelihood of dream content being semantically similar to these intentions.

*This paper is part of the Festschrift in honor of Dr. Robert Stickgold. The open-access publishing fee for this paper was contributed by the Furman University Sleep Laboratory.*

**Key words:** dreams; intentions; prospective memory; natural language processing

### Statement of Significance

We used an objective AI-based large language model analysis to analyze the semantic similarity of dream reports to task plans participants had learned before sleep, and that was either executed before sleep or remained uncompleted and, thus, “active” during the subsequent sleep period. We find that the content of the reported dreams was more similar to the task plans that remained uncompleted than to task plans completed before sleep. Whereas dreams are well known to incorporate past experiences, these findings provide first-time experimental evidence that dreams also incorporate anticipated experiences, i.e. prospective memories for future plans.

Dreams occurring during sleep are known to incorporate memories of past experiences that occurred in a state of wakefulness, with the precise relationship between the waking experience and dream content representing one central question in this research field [1–4]. How are dreams constructed from waking experiences? Dream content appears to be particularly linked to new learning, although this idea is controversially discussed [2–4]. In fact, novel waking-life experiences have been shown to be incorporated into the content of non-rapid eye movement (NonREM) dreams, and this incorporation appeared to be specifically associated with the formation of new memories [5–7] suggesting that dreams reflect memory processing during sleep. Moreover, dream incorporation is enhanced when the task experience is highly engaging, e.g. through computer games [8, 9], virtual environments [10, 11], or stressful situations [12], i.e. all factors known to also enhance learning and the formation of new memories. In light of this link to memory formation, dream content has been proposed to be influenced by the neuronal replay of memory representations during sleep, as a core mechanism underlying memory consolidation [7–10, 13, 14]. Such influence of ongoing memory processing on dream content may extend—beyond the precise replay of cell ensemble firing patterns that occurred during learning—also to the reactivation of networks (semantically) connected to the replaying ensembles during sleep [15–19]. It is, however, important to note that although memory reprocessing can bias the dream content, the dream is rarely a mere replay of an episodic wake experience per se, but rather incorporates individual aspects connected to the waking experience [20].

Memories in the brain do not merely serve to represent the past but to predict the future. A first experimental demonstration of this future-oriented aspect of memory has been provided by Bluma Zeigarnik in 1927 [21], who showed that participants remembered tasks better that were interrupted and that remained to be completed in the future than tasks they had completed. The effect was explained such that a task that has already been started establishes a task-specific “tension” which improves accessibility of the relevant task memories, and which is only relieved upon completion of the task. The Zeigarnik effect has been confirmed and elaborated in numerous recent experiments on the nature of prospective memory for intentions and action plans (e.g. [22, 23]). Importantly, this research indicated that prospective memories for plans strongly benefit from periods of post-encoding sleep [24–28]. Specifically, this benefitting effect of sleep was significantly stronger for action plans that remained uncompleted before the experimental sleep period (due to an experimental interruption of task performance or an instructed delayed execution) than for plans that were completed before the sleep period [29].

Evidently, memory shares its distinct prospective orientation with dreams which have been considered “the brain’s offline efforts to distill projections of the future” [30]. This shared prospective focus might be the factor biasing dream content toward the acutely processed memories. Specifically, we propose that dreams incorporate more elements from task plans made during wakefulness, if the respective task is not yet completed and the intention to complete the plan is kept active across sleep. To test this, participants were given three different task plans before sleep. One plan was completed before the experimental sleep period (completed). Of the remaining two plans, one was never started by the participants before sleep, as they were informed it would need to be completed the next morning (uncompleted). The other plan was started but interrupted before sleep,

with participants instructed to finish it the following morning (interrupted).

Crucially, we used an objective approach to analyze the similarity between the task plans and the dream reports resulting from multiple awakenings during the postacquisition sleep period. Recently, strong efforts have been made to objectively analyze dreams [31–37], with particularly impressive results from neural decoding approaches in which machine-learning models predict the reported dream contents based on measured brain activity [38]. Here, we deployed a large language model to match objectively a dream report with the respective task plans, with the results from this objective analysis confirming that the obtained dream reports were on average more similar to the uncompleted and interrupted task plans than to the completed task plan.

## Methods

### Participants

A total of 20 young, healthy participants aged between 19 and 29 years ( $M = 24.15 \pm 3.08$ ) took part in the study, including 11 female and 9 male participants. All participants were native German speaking, non-smokers, and reported not to suffer from any neurological or psychiatric conditions. They did not take any medications (except oral contraceptives), iron, or vitamin supplements. Individuals reporting difficulties to fall or stay asleep, nightmares, or other sleep disorders were also excluded. All participants reported to follow a regular sleep/wake schedule with >6 hours of sleep per night and no shift work, night duties, or long-distance flights with jet lag in the 6 weeks prior to the experiment. Only participants were included who assessed themselves as individuals “dreaming on occasion” and normally having at least “fairly good recall” of their dreams.

Participants spent one adaptation night in the sleep lab before the experimental night, in order to familiarize them to sleeping with EEG electrodes. Data from one participant were excluded from the analysis because of too poor sleep quality, reducing the sample to a total of  $N = 19$  participants. Before the start of each experimental night, the current well-being and health of the participants were assessed, including whether they had experienced any unusual kind of stress, and had refrained from napping, alcohol, and caffeine on that day before the experimental night. Participants gave written informed consent and were paid for participation. The study was approved by the local ethics committee of the University of Tübingen.

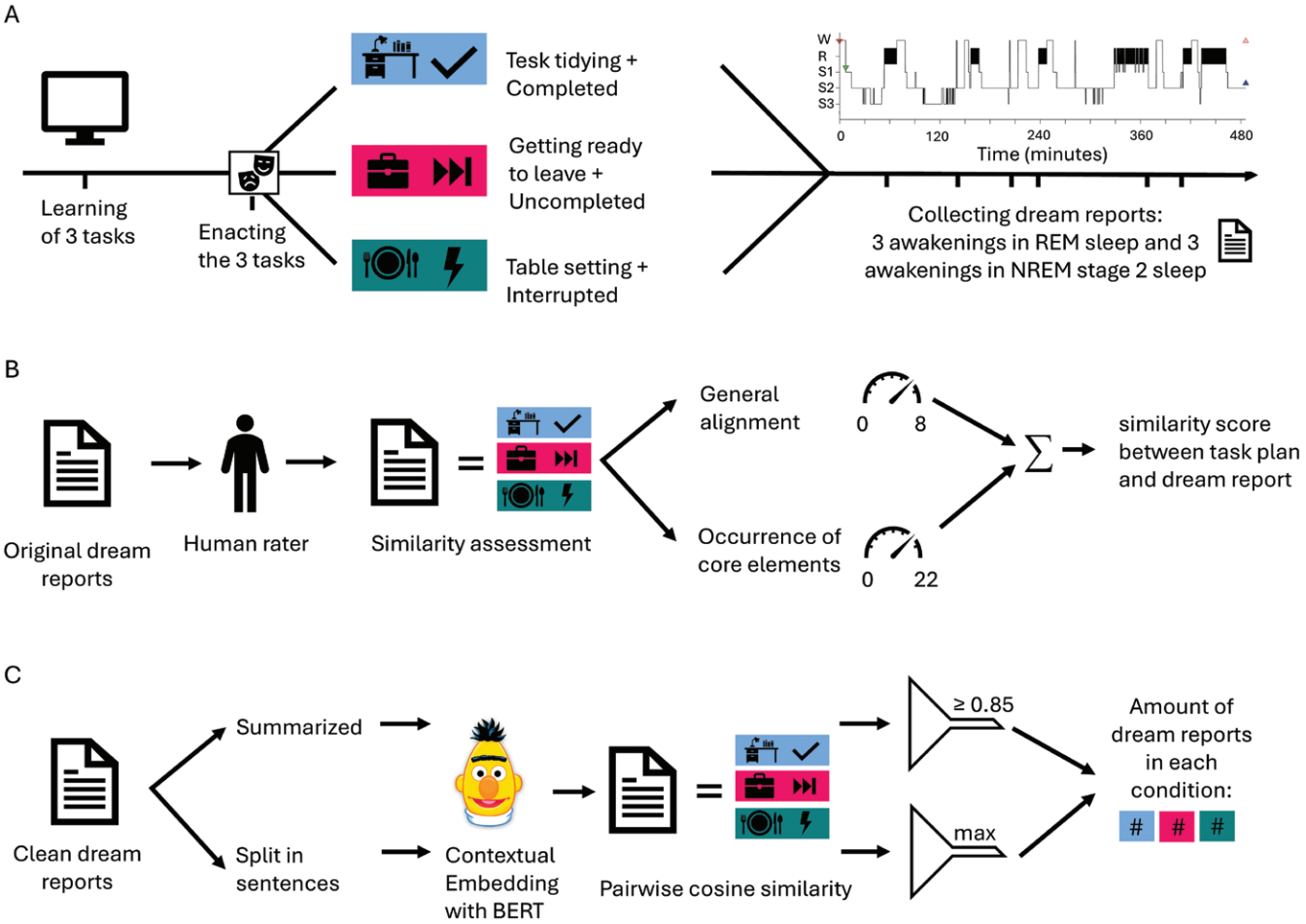
### Design and procedures

The study followed a within-subject design, examining the correspondence of dream reports obtained in a single experimental night to each of three different task plans with different execution status (completed, uncompleted, and interrupted). The task plans were entitled “Setting the table,” “Tidying the desk,” and “Getting ready to leave.” The first two scripts were derived from another study [39], while the third one was newly developed for the purpose of this study. Each task plan was assigned to one of the three execution statuses. Participants learned the scripts for all task plans in the evening and performed them either afterward (for the completed and interrupted conditions) or in the next morning (uncompleted). The assignment of task plans to the execution status conditions was balanced across conditions (completed – C, interrupted – I, uncompleted – U) distributing as follows: “Setting the table” (C = 7; I = 7; U = 6), “Tidying

the desk” (C = 6; I = 6; U = 8), “Getting ready to leave” (I = 7; C = 7; U = 6). Also the order of execution status conditions for the tasks to be performed in the evening was balanced across participants such that the “completed-interrupted” order occurred nine times and the “interrupted-completed” order occurred 11 times. During the nocturnal sleep period, dream reports were gathered during awakenings from either rapid eye movement (REM) sleep or NREM stage 2. The participants were informed that the experiment investigated the effect of sleep on memory for specific task plans and that awakenings for dream reporting would occur.

The experimental procedure is illustrated in Figure 1A. All subjects reported to the laboratory at 8:00 pm. First, a questionnaire regarding participant data was completed to ensure all experimental inclusion criteria were met. Then, EEG electrodes were attached, and the participants completed the Stanford Sleepiness Scale (SSS) and a mood questionnaire, followed by the Regensburg Word Fluency Test (RWT) and a Vigilance Task (VT), both assessing executive cognitive functions.

After performance on these tests, the participants learned the action sequences for the three task plans as described in [41]. The plans were entitled “Desk tidying,” “Getting ready to leave,” and “Setting the table.” Each task plan comprised an action sequence of five subtasks (Supplementary Figure 1). For Desk tidying the subtasks were (1) opening a file, (2) filing documents, (3) sharpening a pencil, (4) sorting index cards, and (5) stacking articles. For Getting ready to leave, the subtasks were (1) shutting down the computer, (2) closing the window, (3) putting on a coat, (4) putting on a backpack, and (5) switching off the light; and for Setting the table, the subtasks were (1) spreading out the tablecloth, (2) distributing tableware, (3) polishing glasses, (4) folding napkins, and (5) lighting candles. The subtasks were presented sequentially on a computer screen, with the title of the task plan displayed above each subtask. Every subtask was shown for 6 seconds in a fixed order. Following this sequential presentation, the task title and all five subtasks were presented together for 30 seconds to allow participants to consolidate the information. This process was repeated three times



**Figure 1.** Experimental procedure and analysis. (A) Experimental procedure for an example participant who completed the task plan “Tidying the desk” before sleep, and was told to execute the task plan “Getting ready to leave” the next morning. On the “Table setting” task plan he/she was interrupted when implementing it before sleep and asked to complete it the next morning. Right: sleep profile for subsequent nighttime, indicated 6 awakenings for oral dream report collection. (B) Analysis of dream reports based on a traditional rating-based analysis of dream reports, where two human raters were asked to rate the similarity of dream reports to the task plans, relying on a standardized scale by Schredl [40]. The similarity score, here, combined points given for the rated overall similarity between a dream report and a specific task plan (0–8 points), and points given for the rated similarity of the reports with core elements of the respective task plans (0–22 points). (C) Analysis of dream reports based on an AI large language model (GermanBERT): dream reports were transcribed into written texts and filler words and grammatical errors were removed. Reports were then either automatically summarized or split into sentences. Semantic similarities between task plan descriptions and dream reports were then determined through contextual embedding and calculating cosine similarities. Summarized dream reports passing the 0.85 criterion threshold and maximum values of dream reports split into sentences, respectively, were counted for each execution status condition.

for each task plan, with one complete cycle through the three plans constituting a learning trial.

After the initial learning phase, an immediate recall test was performed where each task title was presented, and the participant was asked to recall (verbally) the five subtasks of the plan in the correct order. It was ensured that the participants memorized the exact wording and order of the subtasks. If they did not achieve 100% correct recall on all three task plans during two consecutive recall tests, additional learning trials were conducted.

The second (and following) learning trials followed the same procedure as the first, but with each task plan being presented only once. On average, participants required  $2.9 \pm 0.9$  learning trials (range 2–5) to meet the recall criterion. To enhance task involvement, the participants were informed that they would receive an additional 15 € for correctly recalling all three task plans (5 € per plan).

After learning, the participants were informed which two task plans were to be performed in the evening and which one in the morning. Plans were executed under observation and evaluation by the experimenter, with the participant not being allowed to ask any questions while performing the task plans.

The materials used to execute the plans were prepared beforehand. Once the participant had performed on the first plan in the evening he/she left the room so that the materials could be reset for the next task plan. For the interrupted condition, the experimenter interrupted the execution of the task plan after the participant had completed the first subtask. Under the pretext that an error had occurred, the participant was instructed to perform this task the next morning. After performing the two task plans, the participants went to bed.

For the experimental sleep interval, lights were turned off at the participant's habitual bedtime. After about 3 hours of sleep or latest at 2:00 am, the first awakening occurred. It was ensured by visual inspection of the ongoing polysomnographic recordings, that before an awakening the sleep stage was stable for at least 10 minutes. Awakenings were done up to six times per night, from NREM stage 2 and REM sleep to cover both NonREM and REM sleep-associated dreams [42–45]. Subsequent awakenings always were carried out after the participant had regained sleep for at least 30 minutes. For awakening, lights were turned on and the participant was addressed by their name and asked to sit up and put on a headset for voice recording. A standard set of questions followed: (1) Tell me everything that was going through your mind before you were woken up. (2) Can you remember any details? (3) And further? (4) Was it a dream or a thought? and (5) Was it pleasant, unpleasant, or neutral? Question 3 was repeated unless the participant explicitly reported not having any further memory. Lights were turned off for the participant to return to sleep once no further details came into their mind. The participant was awakened the next morning at about his/her usual wake-up time. The SSS, mood questionnaire, RWT, and VT were performed a second time about 30 minutes after awakening.

## Polysomnographic recordings

Throughout the night, the EEG was recorded from six channels (F3, F4, C3, C4, P3, and P4) referenced to two electrodes attached to the mastoids (M1 and M2) using a BrainAmp DC amplifier (BrainProducts, Munich, Germany). The ground electrode was placed on the forehead. Impedances were always kept below 5 kOhm. Additionally, vertical and horizontal eye movements were measured (VEOG and HEOG) as well as an electromyogram (from two electrodes placed on the chin). Signals were band-pass

filtered between 0.3–30 Hz (EEG and EOG signals) and 5–150 Hz (EMG signal), sampled at 500 Hz and stored for offline analyses. Visual scoring of 30-second polysomnographic records followed the criteria outlined by Rechtschaffen and Kales [46].

## Dream report analysis

Audio-recorded dreams were transcribed into written texts before analysis. To quantify the extent to which the task plans at the different execution states were incorporated into the dream reports, we analyzed the semantic similarity between dreams and respective task plans. These similarity analyses were performed in two different ways, first using a traditional approach based on subjective ratings and, second, based on an AI large language model.

For the rating-based approach, two colleagues, sleep experts with no special experience in the assessment of dreams, were asked to rate the dream reports according to a standardized scale derived from a previous report by Schredl [40]. The first part of the scale required to rate the general extent of alignment between the reported dream and the specific task plans between 0 and 8 (indicating no vs. high correspondence). The second part of the scale aimed at a similarity rating based on the occurrence of certain core elements characterizing the task plans. For this, for each task plan, 11 core elements, i.e. objects, and activities that characterized the plan, were selected. The rating required to assign, for each of these core elements, a score between 0 and 2 with 0 indicating that the element did not occur in a dream report, 1 indicating that the element occurred metaphorically or indirectly, and 2 indicating that the element was directly named in the dream report, with the sum of the 11 scores defining the similarity between the dream report and the respective task plans. Raters were blinded as to the execution state conditions and as to whether a dream was reported after a NonREM or REM sleep awakening.

For the second analysis (illustrated in Figure 1B), we used a large language model to objectively quantify the extent of dream incorporation. Dream incorporation was measured by the degree of semantic similarity between the task plans and the dream reports. Semantic representations of the task plans and dream reports were extracted using the transformer-based representational model Bidirectional Encoder Representations from Transformers (BERT) [47], a natural text embedding model capable of quantifying semantic textual similarity [48]. Specifically, we used the German version (GermanBERT) [49] which is pretrained on written German texts, i.e. German Wikipedia articles, German OpenLegalData, and German news articles. The parameters of the model were trained by (1) splitting the input texts, i.e. the dream reports and task plans, into tokens representing semantic units (words, subwords) of the input texts, (2) masking some tokens and feeding back the corrupted sentence (with masked tokens) as input into the model, and (3) asking the model to reconstruct the original tokens.

Since BERT was pretrained on written language, the transcripts of the dream reports were additionally edited to remove filler words and to correct grammatical errors. Similarly, the task plans were prepared for these analyses by transforming the description of subtasks in bullet points into full sentences. Then, each dream report was summarized automatically using a BERT model fine-tuned for text summarization ([https://huggingface.co/mrm8488/bert2bert\\_shared-german-finetuned-summarization](https://huggingface.co/mrm8488/bert2bert_shared-german-finetuned-summarization)) before giving it as input to the GermanBERT. The GermanBERT encoder embeddings (vectors that take into account single text units and respective semantic relationships to other units) were used as a

representation of the dream. While these embeddings are influenced by sentence length and parts of speech, this information is not explicitly encoded in the embeddings. We then encoded each of the five subtasks in each of the three task plans into embeddings following the same procedure (without prior summarization). Cosine similarity between the embeddings was calculated resulting in a similarity score for each dream report with each action of the task plans, for which then the maximum similarity score for each task is chosen. Cosine similarity measures how closely aligned two vectors are, regardless of their magnitude, by calculating the cosine of the angle between them. In our study, we used cosine similarity to quantify the semantic relatedness between dream reports and task descriptions, with scores ranging from  $-1$  (opposite meaning) to  $1$  (identical meaning). Because similarity cosine values for the dream reports fell into a more limited range at the higher end of the scale [50], we thresholded cosine values using only values greater or equal to  $0.85$  to focus the analysis on dreams with high similarity scores. We then compared the number of dream reports with the above criterion similarity scores for a given task plan, between the execution status conditions completed, interrupted, and uncompleted.

In a further sentence-wise analysis, we split the written dream reports into single sentences. Based on GermanBERT, we then calculated the cosine similarity score for each sentence of a dream report and each subtask of a task plan. The resulting maximum similarity value for each dream report was then used to allocate the dream report to one of the three execution status conditions according to a forced choice procedure.

## Statistical analyses

Similarity scores for dream reports were generally analyzed using analyses of variance (ANOVA) including repeated measures factors representing the executive status conditions (completed, interrupted, and uncompleted) and the sleep stage (NREM stage 2, REM) prior to the obtained report, with subsequent post-hoc t-tests used to specify the significance of pairwise comparisons. Pearson's correlation coefficients were used to assess interrater reliability in the analyses based on subjective ratings. Given that similarity scores for the dream reports with the above criterion similarity in the AI language model-based analyses were not uniformly distributed among execution status conditions and task

plans, we focused on the statistical analysis of these scores on nonparametric testing using  $\chi^2$  tests. Prior to testing, the number of dream reports above the criterion in each execution status condition and for each task plan was additionally divided by the total number of dream reports obtained for each execution condition and task plan. The  $\chi^2$  test was also used to detect deviations from equal distributions of dream reports collected in different sleep stages, conditions, and tasks. The level of significance was set to  $p = .05$  for all statistical tests, and  $p = .01$ , in case of directed one-tailed testing of hypotheses.

## Results

### Dream report acquisition

A total of 117 awakenings were performed, equally distributed across REM sleep and NREM stage 2, i.e. 59 (50.4%) awakenings were performed in REM sleep and 58 (49.6%) awakenings in NREM stage 2 sleep ( $p = .926$ ,  $\chi^2$  test). A dream was recalled in 86 (73.5%) of the awakenings. Of these 86 dream reports, 51 (59.3%) were collected after REM sleep awakenings and 35 (40.7%) after NREM stage 2 sleep awakenings, resulting in a trend toward more dream reports collected after REM sleep than NREM stage 2 awakenings ( $p = 0.084$ ,  $\chi^2$  test). Since participants' ability to recall dreams at each awakening differed, the number of dream reports was not evenly distributed among task plans and execution status conditions (Table 1). SSS sleepiness scores averaged (mean  $\pm$  standard deviation)  $3.1 \pm 0.85$  in the evening before sleep, and  $2.5 \pm 0.76$  in the next morning. Performance scores on the RWT averaged  $16.4 \pm 4.81$ , in the evening, and  $16.6 \pm 4.32$  in the next morning.

### Dream incorporation assessed by subjective ratings

The two raters only moderately agreed on their ratings of the dream reports. Although significant, both correlations between their ratings of the general similarity between task plans and dream reports as well as correlations of their judgments based on the occurrence of core elements of the task plans (objects, activities) in the individual dreams were only of medium size ( $0.52 < r < 0.65$ ,  $p < .01$ , Pearson's correlation). ANOVA performed on ratings collapsed across both raters did not reveal any significant difference in similarity ratings between any of the execution

**Table 1.** Total number of dream reports for the task plans and execution status conditions, across all awakenings (top) and separately for awakenings from REM sleep and NREM stage 2 sleep

All reports	Completed	Uncompleted	Interrupted	$\Sigma$
Setting the table	34	25	27	86
Tidying the desk	20	37	29	86
Getting ready to leave	32	24	30	86
$\Sigma$	86	86	86	258
Reports in REM				
Setting the table	19	16	16	51
Tidying the desk	13	21	17	51
Getting ready to leave	19	14	18	51
$\Sigma$	51	51	51	153
Reports in NREM stage 2				
Setting the table	15	9	11	35
Tidying the desk	7	16	12	35
Getting ready to leave	13	10	12	35
$\Sigma$	35	35	35	105

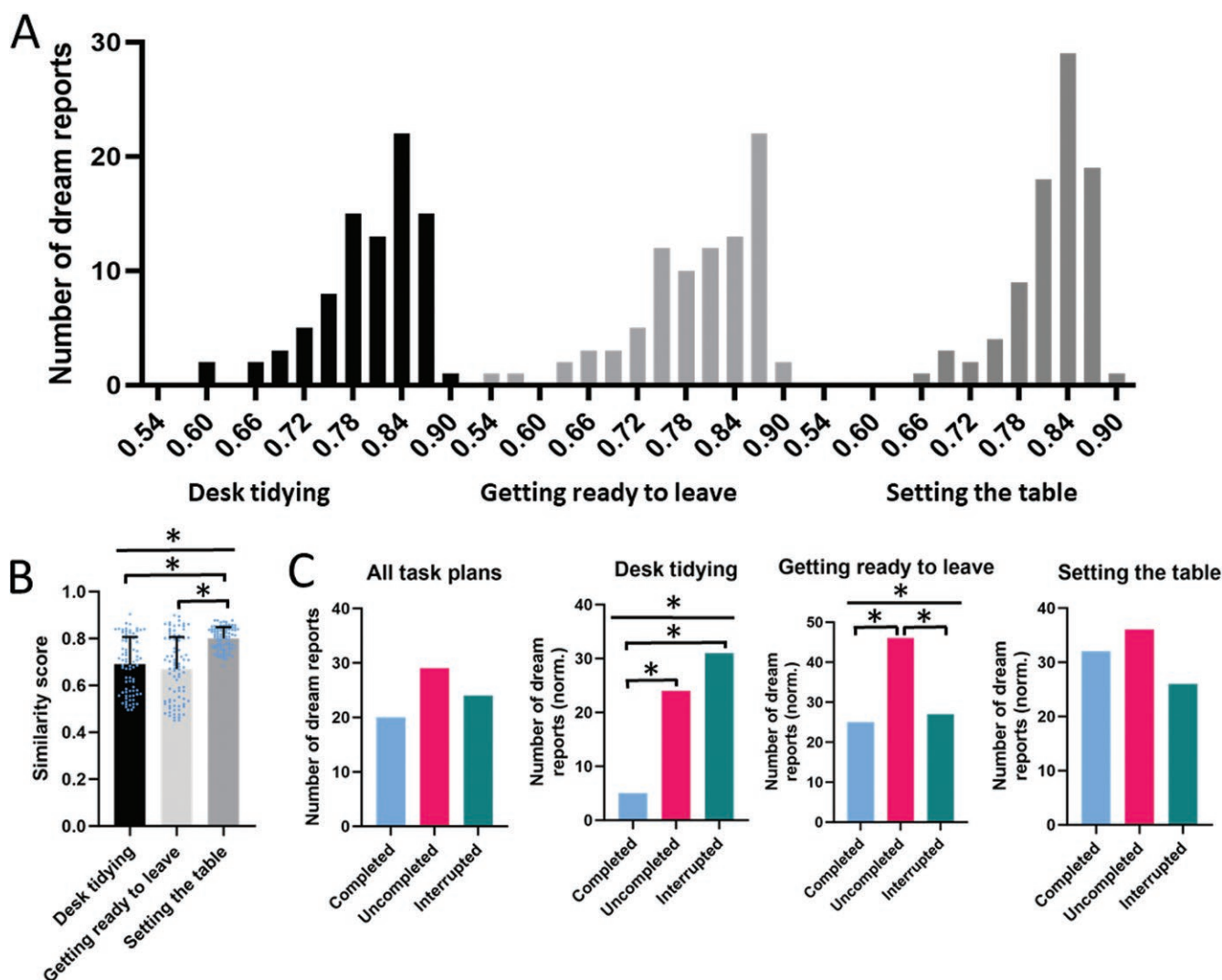
status conditions or awakenings from REM or NREM stage 2 (all  $p > .67$  for respective ANOVA factors). Values collapsed across general similarity ratings and core element-based ratings indicated for one of the rater's highest dream incorporation for the completed task plans ( $1.02 \pm 0.20$ ), medium for uncompleted plans ( $0.98 \pm 0.20$ ), and lowest for interrupted task plans ( $0.80 \pm 0.17$ ) whereas for the other rater, values were highest for the interrupted task plans ( $1.20 \pm 0.22$ ), medium for completed plans ( $0.78 \pm 0.14$ ) and lowest for the uncompleted tasks ( $0.70 \pm 0.15$ ). The overall insufficient agreement between our raters is consistent with a great body of findings in this field of dream content analysis [51, 52] and led us to switch to an AI-based approach.

### Assessment of dream incorporation by an AI-based large language model approach

Here, we determined the semantic similarity of dream reports by applying a large language model (GermanBERT) to transcripts of the reports (Figure 1B). Cosine similarity scores were calculated

for the whole texts, indicating the similarity between a dream report and one of the three task plans. For all task plans, similarity scores ranged between 0.54 and 0.90, with maximum frequencies in the 0.84–0.90 range (Figure 2A), with this distribution indicating that a minimum similarity (of around 0.54) to each of the three task plans is basically reached by any of the dream reports. The direct comparison between task plans, on the other side, indicated that the task plan “Setting the table” yielded distinctly higher similarity scores than the two other task plans ( $F(1.040,88.43) = 116.9$ ,  $p < .0001$ , for ANOVA main effect of task plan), with this effect being independent of the execution status condition the task plan was assigned to ( $p = .49$  for respective interaction effect with execution status factor, Figure 2B). This finding points to an a priori difference in our task materials, with a higher likelihood for the “Setting the table” plan to be similar to a dream report than for the other two plans.

Given that the frequency distribution of similarity scores (Figure 2A) indicated that each dream report shows at least a



**Figure 2.** Whole-text similarity analysis of dream reports. (A) Frequency distribution of dream report similarity scores for each of the three task plans. (B) Mean  $\pm$  SEM similarity scores for the task plans across execution status conditions. Note, that the generally enhanced number of reports with similar to the “Setting the table” task. Asterisks indicate  $p < .05$ , for ANOVA across the three task plans (straight line) and for (unpaired) two-sided  $t$ -tests between task plans (brackets). (C) Number of dream reports with similarity scores  $\geq 0.85$  across all task plans and separately for each task plan. To account for differences in the general similarity scores between the task plans, numbers for the separate plans were normalized by dividing the number of reports with  $\geq 0.85$  similarity with the total number for the respective plan. Asterisks indicate  $p < .05$  two-tailed  $\chi^2$  test, for the comparison across all three execution status conditions (straight line) and for pairwise comparisons between conditions (brackets).

minimum similarity to any of the three task plans (of around 0.54), we focused our analyses on only the dream reports exhibiting substantial similarity to one of the tasks, adopting a criterion similarity score of  $\geq 0.85$ . Indeed, we assumed that higher similarity scores are associated with a higher probability that this similarity was related to specific features of one of the three task plans. Note, although the  $\geq 0.85$  criterion is arbitrary, virtually the same results were obtained with lower criteria, up to  $\geq 0.75$ . Comparing each dream report with each of the three task plans, we revealed that out of all 86 dream reports, 29 reached a semantic similarity score  $\geq 0.85$  to the task plan assigned to the uncompleted execution status condition, 24 dream reports reached the  $\geq 0.85$  criterion for the interrupted tasks, and 20 for the completed task plans. Although descriptively this pattern concurred with our hypothesis of an increased incorporation into dream reports of contents from uncompleted and interrupted task plans in comparison with completed task plans, it did not reach significance ( $\chi^2(2,73) = 1.67, p = .4439$ , for the comparison between execution state conditions), which we hypothetically attributed to the fact that our task plans showed a priori differences in the likelihood of being highly similar to a dream report with the highest likelihood for the “Setting the table” task plan (Figure 2B). Indeed, analyzing separately similarity scores for the different task plans, we found that significantly more dream reports were semantically similar to the “Desk tidying” and “Getting ready to leave” task plans when they were uncompleted (Desk tidying—24, and Getting ready to leave—46) or interrupted (Desk tidying—31, and Getting ready to leave—27) compared to being completed (Desk tidying—5, Getting ready to leave—25; Desk tidying:  $\chi^2(2,60) = 18.10, p < .001$ ; and Getting ready to leave:  $\chi^2(2,98) = 8.23, p < .05$ ; for the comparison across all three execution state conditions, see Figure 2C for pairwise comparisons). On the other hand, for the “Setting the table” task plan no such pattern was obtained ( $\chi^2(2,99) = 2.24, p > .30$ ). We found no comparable differences between the execution status conditions in separate analyses of dream reports obtained after REM sleep awakenings (all  $p > .36$ ) or after NREM stage 2 sleep awakenings (all  $p > .33, \chi^2$  test). Moreover, an exploratory control analysis of word counts for dream reports with the highest similarity to task plans revealed no significant differences between the execution status conditions ( $p = .5052, F(2,28) = 0.7053$ ).

To further validate our large language model-based approach, in a second analysis, we made use of a forced choice method where dream reports were allocated to one of the three execution status conditions after a sentence-wise comparison of dream reports and task plans (Figure 1B). This forced choice approach appeared to be also favorable against the backdrop that the similarity scores of an individual dream report for the three different task plans were generally rather close to each other, i.e. showed relatively low variability in comparisons with the high variability of similarity scores among the different dream reports (averaged across task plans;  $F(1,83) = 0.0014, p < .002$ , for a direct comparison between respective variances). This sentence-wise analysis revealed similarity scores ranging between 0.78 and 0.90 for each of the three task plans, with maximum frequencies around 0.86 (Figure 3A). Similarity scores again significantly differed for the three task plans ( $F(2, 249) = 8.7, p < .0005$ , for ANOVA main effect of task plan), with this effect being independent of the execution status condition the task plan was assigned to ( $F(4, 249) = 0.64, p = .63$ , for ANOVA task plan  $\times$  execution status interaction). When we assigned each dream report to the execution status condition with the highest similarity score for this dream report, we found that the lowest number of dream reports, i.e. 20 reports, were assigned

to the completed condition, the number of assigned reports was intermediate (28 reports) for Uncompleted task plans, and highest for interrupted task plans (38 dream reports ( $\chi^2(2,86) = 5.678, p < .01$ , one-tailed  $\chi^2$  test, for the comparison across all three conditions, Figure 3B and C).

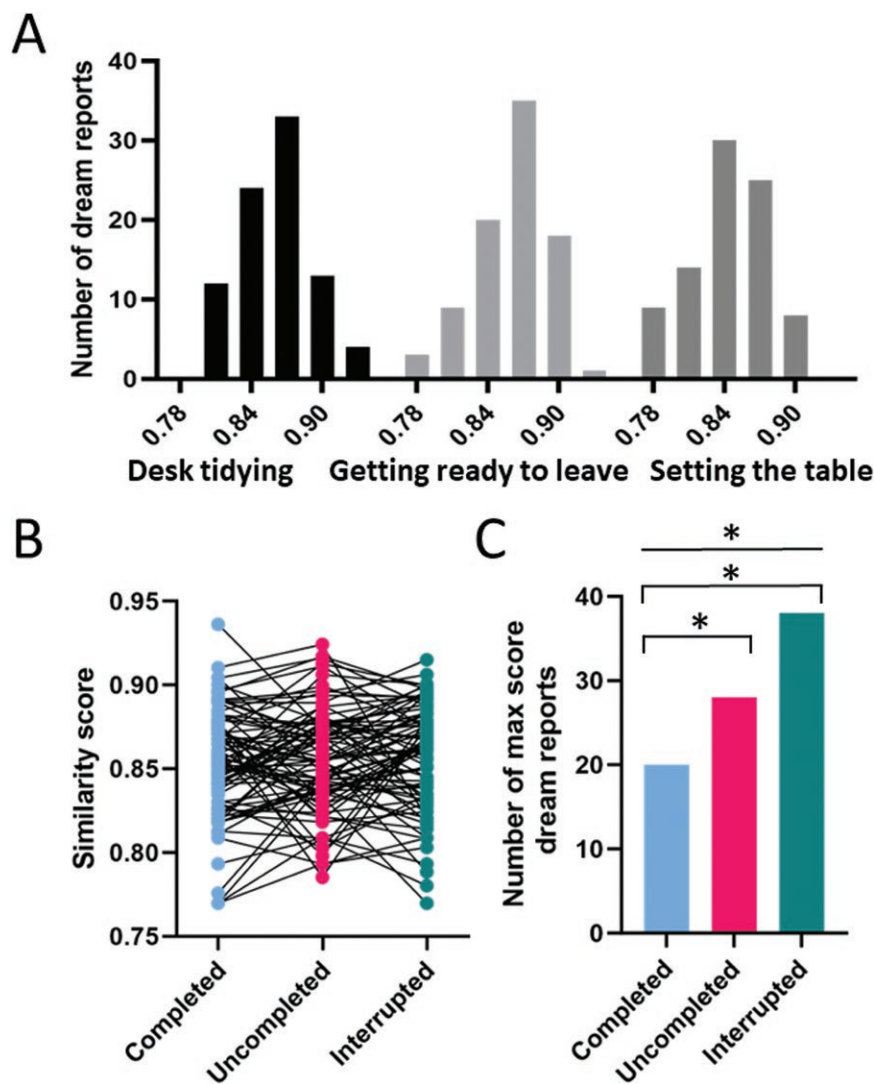
## Discussion

In this study, we explored whether intentions for future actions influence dream content. Employing an AI-based large language model analysis, we show that task plans that have not been completed before sleep and, hence, remain active during sleep, influence the content of a dream to a greater extent than tasks that are completed before sleep. Specifically, tasks whose execution was interrupted before sleep or whose execution was anticipated for the morning after sleep produced dream reports of greater semantic similarity to these tasks than task plans that were completed before sleep. Whereas firm evidence has been accumulated that dreams incorporate past experiences [10, 12, 33, 53], especially if they are emotional [54–56], our findings provide first-time experimental evidence that dreams also incorporate anticipated experiences, i.e. future plans.

In psychological terms, our findings relate to the well-known Zeigarnik effect [21] or the intention–superiority effect [39] which describes the phenomenon that a planned action is better retained in memory or in a heightened state of activation as long as the plan is not executed [57–59]. It is assumed that a “tension” sometimes carrying also an emotional tone [52], drives ongoing processing of memory representations connected to the plan, as long as it is not executed. This tension and associated processing of plan-related representations is not necessarily conscious, and we here provide evidence that it extends into sleep biasing the content of dream reports. This conceptual view is in line with multiple studies showing that sleep promotes problem solving on tasks that remained unsolved before sleep, probably due to a subliminal ongoing processing of the problem [2, 60, 61]. Indeed, the incorporation of experienced content into dreams has likewise been linked to an ongoing reprocessing of respective memory representations that not only supports the consolidation of respective memory but simultaneously, expresses itself in dream reports that are semantically biased toward the reprocessed memory contents [3, 5, 9, 10, 20].

Studies of prospective memories for plans and intentions have indicated a greater benefit for such memories of uncompleted plans from slow wave sleep (SWS) than REM sleep [26]. Hence, assuming a direct link between processes of memory consolidation and dreaming, one might expect that dream reports after awakening from SWS show a greater similarity to the uncompleted task plans than reports after REM sleep awakenings. The present data remain inconclusive in this regard for two reasons. First, rather than in SWS, we awakened the participants in NREM stage 2 sleep during which the reprocessing of the to-be-consolidated memory representations might be less intense than in SWS, though evidence for such difference is mixed [62–64]. Second, our analyses relied on a rather small number of dream reports. The data set, hence, did not provide sufficient statistical power for reliable analyses on subsets separating dream reports between awakenings from NREM stage 2 and REM sleep, considering the size of  $f = 0.48$  (G\*Power version 3.1.9.7) for the effect of the plan execution status on dream report similarity for our analysis across all (NREM stage 2 and REM sleep) dream reports.

Our finding of an incorporation of uncompleted task plans into dream reports appears to be especially noteworthy in



**Figure 3.** Single sentences-based similarity analysis of dream reports. (A) Frequency distribution of dream report similarity scores for each of the three task plans. (B) Similarity scores for each of the 86 dream reports for the task plans assigned to the completed, uncompleted, and interrupted condition in the respective session. Solid lines connect scores for the same dream report. (C) The number of dream reports with maximum similarity score, for each execution status condition. Each dream report was allocated to the execution status condition showing the maximum similarity score for this report.  $p < .05$  one-tailed  $\chi^2$  test, for the comparison across all three conditions (straight line) and for pairwise comparisons between execution status conditions (brackets).

that it derives from an objective large language model-based machine-learning approach, i.e. an approach which, except for a most recent study [65], has so far not been used for dream content analyses. This large language model-based approach overcomes the weakness of traditional analyses of dream reports based on subjective ratings which are notoriously unreliable suffering from modest inter-rater agreements [66, 67]. Here, using two independent raters to classify dream reports according to their similarity to the different task plans, we also found only rather low interrater reliability and, consequently no distinct differences in similarity between dream reports and task plans. Conceivably, we could have strengthened inter-rater reliability by a more intense prior training of our raters on “sham reports” including the discussion of discrepant scores between raters [67]. Perhaps, we could have also enhanced our ratings if—like in other studies [68, 69]—we had additionally asked our participants themselves to rate the similarity of their dream reports with the task plans. Exclusively relying on ratings by other persons, our analyses did not yield any conclusive results.

Nevertheless, although objective, our large language-model-based machine-learning approach bears several limitations, altogether calling for further confirmation of our main findings. First of all, we applied the large language model-based approach post-hoc, and only after the subjective ratings turned out to be unreliable. Related to this, our task plans were not particularly tailored for a large language model-based analysis of their semantic similarities. Basically, the task plans turned out to be too similar to each other, resulting in a large overlap between the task plans with respect to their similarity to the dream reports. While we adopted our task plans from a foregoing study [39] targeting the persistent activation of intentions in memory, task plans with greater semantic differences in the activities, objects, and contexts might have increased the differences in similarity between plans and dream reports. Interestingly, the “Setting the table” plan yielded significantly higher similarity scores than the two other task plans, suggesting that certain activities may be more prone to dream incorporation than others. Obviously, future studies adopting new task plans should rule out such a



priori differences in task plans before experimental use. Such future studies should also overcome other limitations of our study arising, e.g. from a rather crude assessment of sleep lacking occipital recordings which may be particularly important for a precise sleep scoring assessment and therefore relevant in analyses of (visual) dreams.

Our large language model-based approach, using GermanBERT as an embedding extractor, revealed significant differences in dream report similarity that confirmed our a priori hypotheses, supporting the validity of this approach. However, statistical significance per se does not necessarily imply that this approach is also the most valid and optimal. It is to emphasize, however, that we could in principle (internally) replicate our findings here, using two different approaches, i.e. a text-based and a sentence-based, strategy of detecting similarity between dream reports and plans, in combination with a statistical assessment of different target parameters (number of above-criterion similarity reports vs. forced choice allocation of the maximum similarity report). This mutual confirmation further corroborates the validity of our approach, although there may be other more optimal strategies for detecting semantic similarity. An example is BERTScore [50] which is a language generation evaluation metric based on BERT contextual embeddings. In contrast to our approach, in which we compute cosine similarities between embeddings of task plans and dream reports, it computes the similarity on a token level, taking into account their context via contextual embeddings, i.e. a strategy potentially allowing for more fine-grained comparisons between dream reports and task plans. Generally, the use of large language models for analyzing dream content is in its beginnings but, eventually, may turn out a promising tool also for other topics of dream research such as the differentiation of reports of dreams versus more or less emotional wake experiences as well as the differentiation of dream reports among individuals, with potentially important therapeutical implications. Whatever the case, GermanBERT administered to the present data set revealed results confirming our a priori hypotheses. Nevertheless, our approach and findings require further confirmation and external validation, ideally through application to other similar data sets, although other validation strategies are conceivable.

## Supplementary Material

Supplementary material is available at *SLEEP Advances* online.

## Acknowledgments

The authors thank Lisa Bastian for technical assistance in the sleep lab, as well as Ilona Sauter and Astrid Günther for support in subjective ratings of dream reports.

## Funding

This research was supported by grants from the Deutsche Forschungsgemeinschaft (FOR 5434) and the European Research Council (ERC AdG 883098 SleepBalance) to J.B as well as (ERC StG 853489 DEXIM) to Z.A.

## Disclosure Statement

Financial disclosure: None. Non-financial disclosure: The authors declare no conflicts of interest or competing interests related to this work.

## Author Contributions

Julia Fechner (Formal analysis [lead], Writing—original draft [lead]), Maren Born (Investigation [lead], Methodology [equal]), Massimiliano Mancini (Formal analysis [equal], Software [equal]), Zeynep Akata (Methodology [equal], Software [equal]), Philipp Haag (Writing—original draft [equal]), Susanne Diekelmann (Conceptualization [equal], Supervision [equal]), and Jan Born (Conceptualization [lead], Funding acquisition [equal], Methodology [equal], Project administration [equal], Supervision [equal], Writing—original draft [equal])

## Data Availability

Data supporting the findings of this study are available from the corresponding author upon request.

## References

1. Freud S. The Interpretation of Dreams Sigmund Freud. Published online 1900. [https://www.dirzon.com/file/The\\_Interpretation\\_Of\\_Dreams.pdf](https://www.dirzon.com/file/The_Interpretation_Of_Dreams.pdf). Accessed May 29, 2024.
2. Horne JA. REM sleep — by default? *Neurosci Biobehav Rev*. 2000;**24**(8):777–797. doi:10.1016/s0149-7634(00)00037-3
3. Stickgold RS. Off-line memory reprocessing. *Trends Cogn Sci*. 1998;**2**(12):484–492. doi:10.1016/S1364-6613(98)01258-3
4. Vertes RP, Eastman KE. The case against memory consolidation in REM sleep. *Behav Brain Sci*. 2000;**23**(6):867–76. doi:10.1017/s0140525x00004003
5. Wamsley EJ, Stickgold R. Dreaming and offline memory processing. *Curr Biol*. 2010;**20**(23):R1010–R1013. doi:10.1016/j.cub.2010.10.045
6. Oudiette D, Paller KA. Upgrading the sleeping brain with targeted memory reactivation. *Trends Cogn Sci*. 2013;**17**(3):142–149. doi:10.1016/j.tics.2013.01.006
7. Wamsley EJ. Dreaming and offline memory consolidation. *Curr Neurol Neurosci Rep*. 2014;**14**(3):1–7. doi:10.1007/S11910-013-0433-5
8. Stickgold R, Malia A, Maguire D, Roddenberry D, O'Connor M. Replaying the game: hypnagogic images in normals and amnesics. *Sci (New York, N.Y)*. 2000;**290**(5490):350–353. doi:10.1126/science.290.5490.350
9. Wamsley EJ, Perry K, Djonlagic I, Reaven LB, Stickgold R. Cognitive replay of visuomotor learning at sleep onset: temporal dynamics and relationship to task performance. *Sleep*. 2010;**33**(1):59–68. doi:10.1093/sleep/33.1.59
10. Wamsley EJ, Tucker M, Payne JD, Benavides JA, Stickgold R. Dreaming of a learning task is associated with enhanced sleep-dependent memory consolidation. *Curr Biol*. 2010;**20**(9):850–855. doi:10.1016/j.cub.2010.03.027
11. Solomonova E, Stenstrom P, Paquette T, Nielsen T. Different temporal patterns of memory incorporations into dreams for laboratory and virtual reality experiences: relation to dreamed locus of control. *Int. J. Dream Res*. 2015;**8**(1):10–26. doi:10.11588/ijodr.2015.1.16611
12. Malinowski J, Horton CL. Evidence for the preferential incorporation of emotional waking-life experiences into dreams. *Dream*. 2014;**24**(1):18–31. doi:10.1037/a0036017
13. Kusse C, Shaffii-Le Bourdieu A, Schrouff J, Matarazzo L, Maquet P. Experience-dependent induction of hypnagogic images during daytime naps: a combined behavioural and EEG study. *J Sleep Res*. 2012;**21**(1):10–20. doi:10.1111/j.1365-2869.2011.00939.x
14. Klinzing JG, Niethard N, Born J. Mechanisms of systems memory consolidation during sleep. *Nat Neurosci*. 2019;**22**(10):1598–1610. doi:10.1038/s41593-019-0467-3

15. Peigneux P, Laureys S, Fuchs S, et al. Learned material content and acquisition level modulate cerebral reactivation during posttraining rapid-eye-movements sleep. *Neuroimage*. 2003;**20**(1):125–134. doi:[10.1016/s1053-8119\(03\)00278-7](https://doi.org/10.1016/s1053-8119(03)00278-7)
16. Schönauer M, Geisler T, Gais S. Strengthening procedural memories by reactivation in sleep. *J Cogn Neurosci*. 2014;**26**(1):143–153. doi:[10.1162/jocn\\_a\\_00471](https://doi.org/10.1162/jocn_a_00471)
17. Buzsáki G. Two-stage model of memory trace formation: a role for “noisy” brain states. *Neuroscience*. 1989;**31**(3):551–570. doi:[10.1016/0306-4522\(89\)90423-5](https://doi.org/10.1016/0306-4522(89)90423-5)
18. Skaggs WE, McNaughton BL. Replay of neuronal firing sequences in rat hippocampus during sleep following spatial experience. *Sci. (New York, N.Y.)*. 1996;**271**(5257):1870–1873. doi:[10.1126/science.271.5257.1870](https://doi.org/10.1126/science.271.5257.1870)
19. Louie K, Wilson MA. Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron*. 2001;**29**(1):145–156. doi:[10.1016/s0896-6273\(01\)00186-6](https://doi.org/10.1016/s0896-6273(01)00186-6)
20. Stickgold R, Hobson JA, Fosse R, Fosse M. Sleep, learning, and dreams: off-line memory reprocessing. *Sci. (New York, N.Y.)*. 2001;**294**(5544):1052–1057. doi:[10.1126/science.1063530](https://doi.org/10.1126/science.1063530)
21. Zeigarnik B. Über das Behalten von erledigten und unerledigten Handlungen [engl. Translation: on the retention of completed and uncompleted activities]. *Psychol Forsch*. 1927;**9**(2):1–85.
22. Möschl M, Fischer R, Bugg JM, Scullin MK, Goschke T, Walser M. Aftereffects and deactivation of completed prospective memory intentions: a systematic review. *Psychol Bull*. 2020;**146**(3):245–278. doi:[10.1037/bul0000221](https://doi.org/10.1037/bul0000221)
23. Matos P, Albuquerque PB. From retrospective to prospective memory research: a framework for investigating the deactivation of intentions. *Cogn Process*. 2021;**22**(3):411–434. doi:[10.1007/s10339-021-01016-7](https://doi.org/10.1007/s10339-021-01016-7)
24. Diekelmann S, Wilhelm I, Wagner U, Born J. Sleep improves prospective remembering by facilitating spontaneous-associative retrieval processes. *PLoS One*. 2013;**8**(10):e77621. doi:[10.1371/journal.pone.0077621](https://doi.org/10.1371/journal.pone.0077621)
25. Barner C, Seibold M, Born J, Diekelmann S. Consolidation of prospective memory: effects of sleep on completed and reinstated intentions. *Front Psychol*. 2017;**7**:2025. doi:[10.3389/fpsyg.2016.02025](https://doi.org/10.3389/fpsyg.2016.02025)
26. Diekelmann S, Wilhelm I, Wagner U, Born J. Sleep to implement an intention. *Sleep*. 2013;**36**(1):149–153. doi:[10.5665/sleep.2322](https://doi.org/10.5665/sleep.2322)
27. Scullin MK, Gao C, Fillmore P, Roberts RL, Pruet N, Bliwise DL. Rapid eye movement sleep mediates age-related decline in prospective memory consolidation. *Sleep*. 2019;**42**(6). doi:[10.1093/sleep/zsz055](https://doi.org/10.1093/sleep/zsz055)
28. Scullin MK, McDaniel MA. Remembering to execute a goal: sleep on it! *Psychol Sci*. 2010;**21**(7):1028–1035. doi:[10.1177/0956797610373373](https://doi.org/10.1177/0956797610373373)
29. Wilhelm I, Diekelmann S, Molzow I, Ayoub A, Mölle M, Born J. Sleep selectively enhances memory expected to be of future relevance. *J Neurosci*. 2011;**31**(5):1563–1569. doi:[10.1523/JNEUROSCI.3575-10.2011](https://doi.org/10.1523/JNEUROSCI.3575-10.2011)
30. Hancock PA. the royal road to time: how understanding of the evolution of time in the brain addresses memory, dreaming, flow, and other psychological phenomena. *Am J Psychol*. 2015;**128**(1):1–14. doi:[10.5406/amerjpsyc.128.1.0001](https://doi.org/10.5406/amerjpsyc.128.1.0001)
31. Amini R, Sabourin C, De Koninck J. Word associations contribute to machine learning in automatic scoring of degree of emotional tones in dream reports. *Conscious Cogn*. 2011;**20**(4):1570–1576. doi:[10.1016/j.concog.2011.08.003](https://doi.org/10.1016/j.concog.2011.08.003)
32. Wong C, Amini R, De Koninck J. Automatic gender detection of dream reports: a promising approach. *Conscious Cogn*. 2016;**44**:20–28. doi:[10.1016/j.concog.2016.06.004](https://doi.org/10.1016/j.concog.2016.06.004)
33. Fogel SM, Ray LB, Sergeeva V, De Koninck J, Owen AM. A novel approach to dream content analysis reveals links between learning-related dream incorporation and cognitive abilities. *Front Psychol*. 2018;**9**(AUG):387144. doi:[10.3389/fpsyg.2018.01398](https://doi.org/10.3389/fpsyg.2018.01398)
34. Elce V, Handjaras G, Bernardi G. The language of dreams: application of linguistics-based approaches for the automated analysis of dream experiences. *Clocks & Sleep*. 2021;**3**(3):495–514. doi:[10.3390/clockssleep3030035](https://doi.org/10.3390/clockssleep3030035)
35. Fogli A, Maria Aiello L, Quercia D. Our dreams, our selves: automatic analysis of dream reports. *R Soc Open Sci*. 2020;**7**(8):192080. doi:[10.1098/rsos.192080](https://doi.org/10.1098/rsos.192080)
36. Yu CKC. Automated analysis of dream sentiment—the royal road to dream dynamics? *Dream*. 2022;**32**(1):33–51. doi:[10.1037/drm0000189](https://doi.org/10.1037/drm0000189)
37. Zheng X, Schweickert R. Differentiating dreaming and waking reports with automatic text analysis and support vector machines. *Conscious Cogn*. 2023;**107**:103439. doi:[10.1016/j.concog.2022.103439](https://doi.org/10.1016/j.concog.2022.103439)
38. Horikawa T, Tamaki M, Miyawaki Y, Kamitani Y. Neural decoding of visual imagery during sleep. *Sci. (New York, N.Y.)*. 2013;**340**(6132):639–642. doi:[10.1126/science.1234330](https://doi.org/10.1126/science.1234330)
39. Goschke T, Kuhl J. Representation of intentions: persisting activation in memory. *J Exp Psychol Learn Mem Cogn*. 1993;**19**(5):1211–1226. doi:[10.1037//0278-7393.19.5.1211](https://doi.org/10.1037//0278-7393.19.5.1211)
40. Schredl M. Characteristics and contents of dreams. *Int Rev Neurobiol*. 2010;**92**(C):135–154. doi:[10.1016/S0074-7742\(10\)92007-2](https://doi.org/10.1016/S0074-7742(10)92007-2)
41. Barner C, Schmid SR, Diekelmann S. Time-of-day effects on prospective memory. *Behav Brain Res*. 2019;**376**:112179. doi:[10.1016/j.bbr.2019.112179](https://doi.org/10.1016/j.bbr.2019.112179)
42. Aserinsky E, Kleitman N. Regularly occurring periods of eye motility, and concomitant phenomena, during sleep1. *J Neuropsychiatry Clin Neurosci*. 2003;**15**(4):454–455. doi:[10.1176/jnp.15.4.454](https://doi.org/10.1176/jnp.15.4.454)
43. Dement W, Kleitman N. The relation of eye movements during sleep to dream activity: an objective method for the study of dreaming. *J Exp Psychol*. 1957;**53**(5):339–346. doi:[10.1037/h0048189](https://doi.org/10.1037/h0048189)
44. Vogel GW, Barrowclough B, Giesler DD. Limited discriminability of REM and sleep onset reports and its psychiatric implications. *Arch Gen Psychiatry*. 1972;**26**(5):449–455. doi:[10.1001/archpsyc.1972.01750230059012](https://doi.org/10.1001/archpsyc.1972.01750230059012)
45. Oudiette D, Dealberto MJ, Uguccioni G, et al. Dreaming without REM sleep. *Conscious Cogn*. 2012;**21**(3):1129–1140. doi:[10.1016/j.concog.2012.04.010](https://doi.org/10.1016/j.concog.2012.04.010)
46. Rechtschaffen A, Kales A. *A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects*. Bethesda, MD: US Department of Health, Education, and Welfare – National Institutes of Health; 1968.
47. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: human Language Technologies - Proceedings of the Conference. 2018;**1**:4171–4186. <https://arxiv.org/abs/1810.04805v2>. Accessed May 22, 2024.
48. Muennighoff N, Tazi N, Magne L, Reimers N. MTEB: massive text embedding benchmark. *arXiv preprint arXiv:2210.07316*. 2022. doi:[10.48550/arXiv.2210.07316](https://doi.org/10.48550/arXiv.2210.07316)

49. Scheible R, Thomczyk F, Tippmann P, Jaravine V, Boeker M. GottBERT: a pure German Language Model. Published online December 3, 2020. Accessed June 10, 2024. <https://arxiv.org/abs/2012.02110v1>
50. Zhang T, Kishore V, Wu F, Weinberger KQ, Artzi Y. BERTScore: evaluating text generation with BERT. *8th International Conference on Learning Representations, ICLR 2020*. Published online April 21, 2019. <https://arxiv.org/abs/1904.09675v3>. Accessed May 22, 2024.
51. Nielsen TA, Powell RA. The day-residue and dream-lag effects: a literature review and limited replication of two temporal effects in dream formation. *Dream*. 1992;**2**(2):67–77. doi:10.1037/h0094348
52. Nielsen TA, Powell RA. The “dream-lag” effect: a 6-day temporal delay in dream content incorporation. *Psychiatr J Univ Ott*. 1989;**14**(4):561–565. <https://europepmc.org/article/med/2813638>
53. Darsaud A, Wagner U, Baiteau E, et al. Neural precursors of delayed insight. *J Cogn Neurosci*. 2011;**23**(8):1900–1910. doi:10.1162/jocn.2010.21550
54. Malinowski JE, Horton CL. Dreams reflect nocturnal cognitive processes: early-night dreams are more continuous with waking life, and late-night dreams are more emotional and hyperassociative. *Conscious Cogn*. 2021;**88**:103071. doi:10.1016/j.concog.2020.103071
55. Schredl M, Doll E. Emotions in diary dreams. *Conscious Cogn*. 1998;**7**(4):634–646. doi:10.1006/ccog.1998.0356
56. Merritt JM, Stickgold R, Pace-Schott E, Williams J, Hobson JA. Emotion profiles in the dreams of men and women. *Conscious Cogn*. 1994;**3**(1):46–60. doi:10.1006/ccog.1994.1004
57. Szpunar KK, McDermott KB, Roediger HL. Expectation of a final cumulative test enhances long-term retention. *Mem Cognit*. 2007;**35**(5):1007–1013. doi:10.3758/bf03193473
58. Marsh RL, Hicks JL, Bink ML. Activation of completed, uncompleted, and partially completed intentions. *J Exp Psychol Learn Mem Cogn*. 1998;**24**(2):350–361. doi:10.1037//0278-7393.24.2.350
59. Shimizu H. Rehearsal strategies, test expectancy, and memory monitoring in free recall. *Memory*. 1996;**4**(3):265–287. doi:10.1080/096582196388951
60. Wagner U, Gais S, Haider H, Verleger R, Born J. Sleep inspires insight. *Nature*. 2004;**427**(6972):352–355. doi:10.1038/nature02223
61. Verleger R, Rose M, Wagner U, Yordanova J, Kolev V. Insights into sleep’s role for insight: studies with the number reduction task. *Adv Cogn Psychol*. 2013;**9**(4):160–172. doi:10.2478/v10053-008-0143-8
62. Brodt S, Inostroza M, Niethard N, Born J. Sleep—a brain-state serving systems memory consolidation. *Neuron*. 2023;**111**(7):1050–1075. doi:10.1016/j.neuron.2023.03.005
63. Carbone J, Bibian C, Born J, Forcato C, Diekelmann S. Comparing targeted memory reactivation during slow wave sleep and sleep stage 2. *Sci Rep*. 2024;**14**(9057):9057. doi:10.1038/s41598-024-59696-y
64. Schönauer M, Alizadeh S, Jamalabadi H, Abraham A, Pawlizki A, Gais S. Decoding material-specific memory reprocessing during sleep in humans. *Nat Commun*. 2017;**8**(1):15404. doi:10.1038/ncomms15404
65. Bertolini L, Elce V, Michalak A, Widhoezl HS, Bernardi G, Weeds J. Automatic annotation of dream report’s emotional content with large language models. In: *Proceedings of the 9th Workshop on Computational Linguistics and Clinical Psychology (CLPsych 2024)*. St. Julians, Malta: Association for Computational Linguistics; 2024: 92–107.
66. Schredl M, Funkhouser AT, Cornu CM, Hirsbrunner HP, Bahro M. Reliability in dream research: a methodological note. *Conscious Cogn*. 2001;**10**(4):496–502. doi:10.1006/ccog.2001.0522
67. Schredl M, Burchert N, Gabatin Y. The effect of training on inter-rater reliability in dream content analysis. *Sleep and Hypnosis*. 2004;**6**(3):139.
68. Eichenlaub JB, van Rijn E, Phelan M, et al. The nature of delayed dream incorporation (‘dream-lag effect’): personally significant events persist, but not major daily activities or concerns. *J Sleep Res*. 2019;**28**(1):e12697.
69. Van Rijn E, Eichenlaub JB, Lewis PA, et al. The dream-lag effect: selective processing of personally significant events during rapid eye movement sleep, but not during slow wave sleep. *Neurobiol Learn Mem*. 2015;**122**:98–109. doi:10.1016/j.nlm.2015.01.009